

# La relevancia de las métricas de movimiento en la evaluación objetiva de habilidades motoras en cirugía laparoscópica

I. Oropesa García<sup>1,2</sup>, M.K. Chmarra<sup>3</sup>, P. Sánchez-González<sup>1,2</sup>, P. Lamata de la Orden<sup>1</sup>, F.M. Sánchez-Margallo<sup>4</sup>, F.W. Jansen<sup>3,5</sup>, J. Dankelman<sup>3</sup>, E.J. Gómez Aguilera<sup>1,2</sup>

<sup>1</sup> Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, España; {ioropesa,psanchez,lamata,egomez}@gbt.tfo.upm.es

<sup>2</sup> Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Zaragoza, España

<sup>3</sup> Department of BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering (3mE), Delft University of Technology, Delft, Países Bajos; {m.k.chmarra,j.dankelman}@tudelft.nl

<sup>4</sup> Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España, msanchez@ccmijesususon.com

<sup>5</sup> Department of Gynecology, Leiden University Medical Center, Leiden, Países Bajos, f.w.jansen@lumc.nl

## Resumen

*El presente trabajo investiga la relevancia de las métricas de movimiento en la evaluación objetiva de habilidades laparoscópicas, así como su correlación directa a dichas habilidades. Se ha realizado una validación de construcción de tres nuevas tareas para la evaluación de cirujanos. Para registrar los movimientos, se ha empleado el sistema de tracking TrEndo. 42 voluntarios (16 novatos, 22 residentes y 4 expertos) participaron en el experimento. Los resultados muestran que tiempo, camino recorrido y profundidad son métricas de evaluación válidas para un amplio espectro de habilidades. Otras métricas como la brusquedad de movimientos o el tiempo muerto demuestran validez en tareas bi-manuales. Métricas como la brusquedad o la velocidad media muestran un alto grado de independencia con respecto a los objetivos de la tarea a realizar. Se verifica la utilidad de este tipo de métricas, si bien son necesarios nuevos estudios que corroboren los resultados alcanzados.*

## 1. Introducción

La importancia que ha cobrado la cirugía laparoscópica en las últimas décadas se ve reflejada en el continuo aumento de procedimientos que se abordan mediante este tipo de técnicas quirúrgicas. Las múltiples ventajas que suponen para el paciente (menor trauma tisular, menor morbilidad, menor estancia hospitalaria y recuperación más rápida) justifican el cambio en el entorno quirúrgico, donde se imponen ciertas limitaciones al cirujano debido a la manipulación indirecta de los órganos a través de un monitor bidimensional que muestra la señal captada por el endoscopio [1].

La formación y acreditación de nuevos cirujanos versados en este tipo de técnicas está siendo objeto de un cambio de paradigma en los últimos años, de un modelo basado en tutelados maestro-aprendiz hacia programas formativos objetivos y estructurados [2][3]. Estos modelos abogan por retrasar el momento en que los residentes se involucren activamente en una cirugía real hasta el momento en que hayan adquirido las habilidades motoras necesarias. En consecuencia, las primeras etapas de

formación se llevan a cabo en entornos de laboratorio, mediante el uso de simuladores físicos y virtuales [2].

El desarrollo de sistemas de seguimiento de objetos junto con los avances en tecnología software, han permitido refinar los métodos de entrenamiento y evaluación empleados. Entre ellos, podemos encontrar (1) simuladores físicos aumentados con tracking del instrumental quirúrgico, como ICSAD [4] o TrEndo [5]; y (2) simuladores de realidad virtual como LapMentor [6], o SINERGIA [7]. Generalmente, los sistemas de seguimiento emplean tecnologías de sensores, que modifican la ergonomía del instrumental y los hacen poco prácticos para entornos de quirófano [8]. Como alternativa, se propone la posibilidad de construir un sistema basado en la localización espacial del instrumental laparoscópico mediante el análisis del vídeo del endoscopio [9].

Ligados a estos sistemas, existen una serie de métricas de evaluación recurrentes relacionadas con el movimiento del instrumental, que es necesario conocer bien de cara al diseño de un sistema de evaluación basado en vídeo. Sin embargo, en general la validez de construcción (la capacidad de la métrica de distinguir entre cirujanos de distinta experiencia) de muchas de ellas no está claramente establecida [10]. Además, estas métricas se suelen considerar independientemente de la tarea que se lleva a cabo y la habilidad que se evalúa. En consecuencia, la evaluación de la pericia se supedita a la elección de parámetros ofrecidos por el sistema empleado, independientemente de su utilidad real [11].

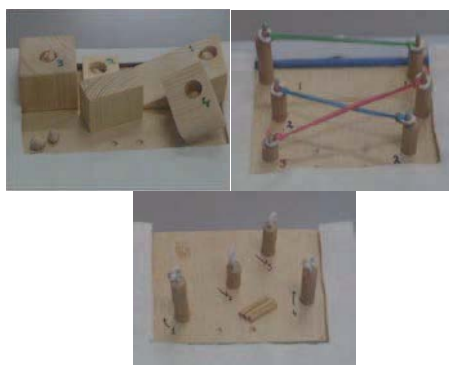
El propósito de este estudio es profundizar en la relevancia de las métricas de evaluación relacionadas con el movimiento del instrumental quirúrgico. En concreto, se pretenden establecer relaciones entre métricas válidas y habilidades quirúrgicas concretas. Para tal fin, se han diseñado y construido tres tareas específicas de simulador físico para la evaluación de habilidades quirúrgicas, en colaboración con expertos docentes de cirugía laparoscópica. Posteriormente, se plantea un experimento de validación combinando seguimiento de instrumental quirúrgico y análisis estadístico de los resultados.

## 2. Materiales y métodos

### 2.1. Tareas empleadas

Las tres tareas de evaluación diseñadas para este experimento se han realizado con el propósito de cubrir y combinar las habilidades motoras básicas en el aprendizaje de un cirujano [1] (Fig. 1):

1. *Grasp and Place* (GP): Ejercicio para realizar con la mano dominante. El objetivo consiste en recoger y colocar secuencialmente en sus orificios una serie de objetos esféricos (garbanzos). El escenario consta de bloques a distintas alturas y orientaciones, así como zonas de oclusión visuales. Las habilidades evaluadas son la coordinación ojo-mano, la percepción espacial y el prensado.
2. *Coordinated Pulling* (CP): El objetivo del ejercicio es la manipulación bi-manual de tres bandas elásticas con el fin de fijar cada una en sus respectivos postes. Cada pareja de postes difiere en altura y orientación respecto a la cámara. Las habilidades evaluadas son la coordinación ojo-mano, la coordinación bi-manual, el prensado y la percepción háptica de fuerzas.
3. *Grasp and Transfer* (GT): En esta tarea, el cirujano debe transferir un objeto cilíndrico a lo largo de un circuito de anillas a diferentes alturas y ángulos haciendo uso de dos pinzas laparoscópicas. Las habilidades evaluadas son la coordinación bi-manual, la percepción espacial, el prensado y algunas habilidades relacionadas con la sutura.



**Figura 1.** Tareas empleadas: Superior, izda: *Grasp and Place*; deha: *Coordinated Pulling*. Inferior: *Grasp and Transfer*

### 2.2. Instalaciones y participantes

El experimento se llevó a cabo en los laboratorios de aprendizaje del Leiden University Medical Centre (LUMC, Leiden, Países Bajos). Un total de 42 participantes realizaron las tres tareas. Al no existir un estándar de clasificación predeterminado, se determinaron cuatro categorías en base a la experiencia: 16 noveles sin experiencia laparoscópica previa (N); 22 residentes, de los cuales 8 constaban de experiencia laparoscópica baja (RB, menos de 10 intervenciones realizadas) y 14 alta (RA, 11-100 intervenciones realizadas; y 4 expertos (E, más de 100 intervenciones realizadas). En una categorización superior, se considera a los N y RB como no expertos (NEx) y a los RA y E como expertos (Ex). Tres de los participantes eran zurdos.

### 2.3. Validación de construcción

Todos los participantes realizaron las tres tareas en el mismo orden. Puesto que su propósito era la validación como elemento de evaluación de habilidades, no de entrenamiento, a los participantes no se les permitió practicar antes del ensayo. Previo a su realización, cada uno recibió una breve explicación del objetivo de cada tarea, de tal forma que cada uno debía razonar la mejor estrategia para abordarlas.

Para asegurar la igualdad de condiciones, la posición de las tareas, la cámara y los puntos de incisión de la cámara e instrumentos se mantuvieron fijos para todos los participantes. Además, se prefijaron para cada tarea las posiciones iniciales y finales del instrumental quirúrgico y el orden de las maniobras a llevar a cabo.

Para el seguimiento del instrumental quirúrgico se empleó el sistema TrEndo (TUDELFT, Delft, Países Bajos). Este sistema permite medir en tiempo real los movimientos del instrumental quirúrgico en entornos de entrenamiento, mediante tres sensores ópticos [5].

Las métricas analizadas en este estudio han sido las siguientes [5][11]:

- Tiempo: Tiempo empleado para completar una tarea (s).
- Camino recorrido: Camino total recorrido por el instrumental laparoscópico, medido en la punta (m).
- Profundidad: Camino total recorrido en la dirección del eje del instrumental (m).
- Velocidad media: Tasa de cambio de la posición del instrumental por segundo. (mm/s).
- Tiempo muerto: Porcentaje de tiempo en el que el instrumental quirúrgico se mantiene parado. Se estableció un umbral de velocidad menor o igual a 5mm/s. (%).
- Brusquedad de movimientos: Cambios abruptos en la aceleración, que reflejan movimientos bruscos del instrumental ( $m/s^3$ ).
- Área angular: Área entre las posiciones más lejanas ocupadas por la punta del instrumental al realizar una tarea ( $rad^2$ ).
- Volumen angular: Volumen de un elipsoide 3D en función de la desviación estándar de los movimientos del instrumental (m).
- Área de superficie: Área máxima abarcada por el instrumental en el plano definido por la superficie de la tarea, calculada como:  $[Max(x) - Min(x)] * [Max(z) - Min(z)]$  ( $cm^2$ ).
- Volumen total: Volumen máximo cubierto por el instrumental dentro del simulador físico, calculado como:  $Área\ de\ superficie * [Max(y) - Min(y)]$  ( $cm^3$ ).

### 2.4. Análisis estadístico

El análisis estadístico se ha llevado a cabo mediante SPSS v. 17 (SPSS Inc., Chicago, IL, U.S.A.). Para establecer la validación de construcción de una métrica, es preciso que ésta muestre diferencias significativas en sus resultados

para los distintos grupos de experiencia considerados. Para tal fin se ha empleado el test de Mann-Whitney, considerando que existe significancia a  $P < 0.05$ .

Para analizar la dependencia de las métricas a las tareas empleadas, se realizó el test de Cronbach para elementos estandarizados. Un valor alto para este test ( $\alpha > 0.7$ ) implica que la métrica es independiente de los objetivos de la tarea, y por extensión, de las habilidades evaluadas.

### 3. Resultados

Los resultados de significancia e independencia se resumen en la Tabla 1. Sólo dos participantes no fueron capaces de terminar alguna tarea: un RB no completó la tarea CP, y un N la tarea GT. Puesto que el objetivo es el análisis de las métricas de movimiento (no directamente relacionados con el resultado final de la tarea) se decidió incluir sus métricas en el análisis.

#### 3.1. Significancia de las métricas

La tarea GP presenta diferencias significativas entre expertos y no expertos para tiempo, camino recorrido, profundidad, velocidad media y tiempo muerto. En general, las diferencias entre grupos se dan sobre todo entre noveles y el resto de los grupos.

CP presenta diferencias significativas para tiempo, camino recorrido, profundidad, velocidad media y brusquedad de movimientos. Entre los distintos grupos predominan las diferencias entre noveles y (a) residentes con alta experiencia, (b) expertos.

GT presenta las mayores diferencias significativas para la mano no dominante (velocidad media, tiempo muerto, brusquedad de movimientos, área angular, área de superficie). Las diferencias entre grupos son mayores que para el resto de tareas, y se encuentran diferencias entre residentes de baja y alta experiencia, así como entre residentes de alta experiencia y expertos.

Métricas	Significancia (Ex - NEx)			Significancia (N - RB - RA - E)			Independencia
	1.GP	2.CP	3.GT	1.GP	2.CP	3.GT	
Tiempo	x	x	x	N-RB, N-RA, N-E	N-RA, N-E, RB-RA	N-E, RB-E, RA-E	0.675
<b>Mano Dominante</b>	<b>1.GP</b>	<b>2.CP</b>	<b>3.GT</b>	<b>1.GP</b>	<b>2.CP</b>	<b>3.GT</b>	
Camino recorrido	x	x		N-RA, N-E	N-RA, N-E	N-E, RA-E	0.638
Profundidad	x	x		N-E	N-RA	N-E, RA-E	0.542
Velocidad media	x	x	x	N-RA	N-RA		0.738
Tiempo muerto	x			N-RA			0.254
Brusquedad							0.738
Área angular				N-RB, N-RA		N-E, RB-E, RA-E	0.269
Volumen angular						N-E	0.071
Área de superficie				N-RB			0.423
Volumen total							0.448
<b>Mano No Dominante</b>	<b>-</b>	<b>2.CP</b>	<b>3.GT</b>	<b>-</b>	<b>2.CP</b>	<b>3.GT</b>	
Camino recorrido		x			N-RA	N-E, RA-E	0.647
Profundidad		x			N-RA	N-E, RA-E	0.517
Velocidad media		x	x		N-RA, N-E	N-RB, N-RA, N-E, RB-E	0.768
Tiempo muerto						N-RB, N-E	0.42
Brusquedad		x	x			N-RA, RB-RA	0.802
Área angular			x				0.38
Volumen angular							0
Área de superficie			x			N-RB	0.174
Volumen total						N-RB	0.321

**Tabla 1.** Significancia e independencia de las métricas para distintos grupos y tareas. El análisis se realiza por separado para la mano dominante y la no dominante. La significancia entre Ex y NEx (2 grupos) se señalan mediante una x. La significancia entre N, RB, RA y E (4 grupos) se expresa por parejas de grupos

#### 3.2. Independencia de las métricas

El análisis de la dependencia revela que área angular, volumen angular, área de superficie y tiempo muerto presentan bajos valores de  $\alpha$ , por lo que se pueden

considerar más dependientes de cada tarea que el resto. Sólo brusquedad de movimientos y velocidad media para la mano no dominante cumplen la condición de independencia. Otras, como tiempo, camino recorrido o velocidad media para la mano dominante presentan

valores cercanos al umbral, y por tanto se puede establecer un grado considerable de independencia frente a las tareas y habilidades evaluadas.

#### 4. Discusión y conclusiones

El presente trabajo resalta la importancia de las métricas de movimiento en los actuales entornos de formación en cirugía laparoscópica, aumentados mediante tracking de instrumental quirúrgico. La validación de las tareas emprendida ha permitido, por un lado, confirmar algunas de las hipótesis ya planteadas en la literatura acerca de la validez general de métricas, como el tiempo o el camino recorrido; y por otro, plantear nuevas vías de investigación en el campo de las métricas.

La principal motivación de este trabajo viene propiciada a partir de la metodología para la evaluación objetiva de cirujanos propuesta en [12], que divide el problema en tres fases: clínico, tecnológico y analítico. Este trabajo se corresponde con la búsqueda de métricas expuesta en el problema clínico de dicha metodología. La búsqueda se ha enfocado en las métricas de movimiento de cara a explorar, en el problema tecnológico, la posibilidad de realizar el seguimiento del instrumental quirúrgico a partir del análisis del vídeo laparoscópico [9].

En un estudio previo, se realizó una correlación entre métricas de evaluación y habilidades entrenadas en cirugía laparoscópica, basada en la literatura de validación [8]. En él, se observaba como en general, métricas como tiempo, camino recorrido o profundidad, presentan un alto grado de validez e independencia. Esto ha sido verificado en el presente trabajo, tal y como muestran los resultados. Otra métrica como es la brusquedad de movimientos también muestra buenos resultados para tareas bimanuales, en especial sobre la mano no dominante. De especial interés resulta el análisis de la velocidad media. Esta métrica, poco explotada en los sistemas actuales [8], muestra resultados significativos en todas las tareas, así como una alta independencia sobre ellas.

El tiempo muerto muestra diferencias significativas en las tareas GP y GT. Puesto que éstas requieren de una aproximación táctica al problema de transportar un objeto, es posible que la métrica refleje hasta cierto punto el tiempo de planificación de estrategia requerido por el cirujano. Los resultados respecto a esta métrica son muy preliminares, y deben ser verificados en futuros estudios.

Finalmente, el estudio refleja las importantes diferencias entre la mano dominante y la no dominante. En general, el análisis de las métricas para la mano no dominante muestra mayores diferencias entre los grupos considerados, además de una consistencia interna mayor.

Es importante no obstante, verificar en detalle las conclusiones obtenidas en este trabajo mediante nuevos estudios. Además, es necesario que el proceso de evaluación global no se base sólo en este tipo de métricas, sino que tenga en cuenta tanto el uso de parámetros de calidad (resultado final de la tarea, errores cometidos), como un análisis de las habilidades de juicio del alumno por parte del evaluador.

#### Agradecimientos

Este trabajo de investigación ha sido financiado por el programa FPU del Ministerio de Ciencia e Innovación. Los autores agradecen su colaboración a todos los participantes del LUMC y del CCMIJU.

#### Referencias

- [1] Usón J, Sánchez-Margallo FM, Pascual S, Climent S. Formación en Cirugía Laparoscópica Paso a Paso. 4th ed. Minimally Invasive Surgery Centre Jesús Usón eds, 2010.
- [2] Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. *British Journal of Surgery*, vol 91, 2004, pp 1549-58.
- [3] Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *British Medical Journal*, vol 1, 1975, pp 447-51.
- [4] Datta V, Mackay S, Mandalia M, Darzi A. The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *Journal of the American College of Surgeons*, vol 193, 2001; pp 479-85.
- [5] Chmarra MK, Klein S, de Winter JCF, Jansen FW, Dankelman J. Objective classification of residents based on their psychomotor laparoscopic skills. *Surgical Endoscopy* vol 24, 2010, pp 1031-1039.
- [6] LAP Mentor laparoscopic surgery simulator for general surgery, gynaecology and urology [http://www.simbionix.com/LAP\\_Mentor.html](http://www.simbionix.com/LAP_Mentor.html) (Consultada: Julio 2011).
- [7] Lamata P, Gómez EJ, Sánchez-Margallo FM, López Ó, Monserrat C, García V, Alberola C, Rodríguez Florido MÁ, Ruiz J, Usón J. SINERGIA laparoscopic virtual reality simulator: Didactic design and technical development. *Computer Methods and Programs in Biomedicine*, vol 85, 2007, pp 273-283.
- [8] Oropesa I, Sánchez-González P, Lamata P, Chmarra MK, Pagador JB, Sánchez-Margallo JA, Sánchez-Margallo FM, Gómez EJ. Methods and tools for objective assessment of psychomotor skills in laparoscopic surgery. *Journal of Surgical Research*, 2011 (In press).
- [9] Sánchez-González P, Cano AM, Oropesa I, Sánchez-Margallo FM, del Pozo F, Lamata P, Gómez EJ. Laparoscopic Video Analysis for Training and Image Guided Surgery. *Minimally Invasive Therapies and Allied Technologies*, 2010 (In press).
- [10] van Hove PD, Tuijthof GJM, Verdaasdonk EGG, Stassen LPS, Dankelman J. Objective assessment of technical surgical skills. *British Journal of Surgery*, vol 97, 2010, pp 972-987.
- [11] Cotin S, Stylopoulos N, Ottensmeyer M, et al. Metrics for Laparoscopic Skills Trainers: The Weakest Link! Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002. Berlin / Heidelberg: Springer, 2002, pp 35-43.
- [12] Oropesa I, Sánchez-González P, Cano AM, Lamata P, Sánchez-Margallo FM, Gómez EJ. Objective evaluation methodology for surgical motor skills assessment. *Minimally Invasive Therapies and Allied Technologies*, vol 10, 2010, pp 55-56.